

(19) 世界知的所有権機関  
国際事務局



(43) 国際公開日  
2002 年10 月3 日 (03.10.2002)

PCT

(10) 国際公開番号  
WO 02/077867 A1

- (51) 国際特許分類: G06F 17/30, 15/00
- (21) 国際出願番号: PCT/JP02/02639
- (22) 国際出願日: 2002 年3 月20 日 (20.03.2002)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:  
特願2001-90700 2001 年3 月27 日 (27.03.2001) JP
- (71) 出願人 (米国を除く全ての指定国について): 三菱  
スペース・ソフトウェア株式会社 (MITSUBISHI  
SPACE SOFTWARE CO., LTD.) [JP/JP]; 〒105-6137 東  
京都 港区 浜松町二丁目 4 番 1 号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 末木 源一郎

(SUEKI, Genichiro) [JP/JP]; 〒247-0065 神奈川県 鎌倉市 上町屋 5 2 4 番地 三菱スペース・ソフトウエア株式会社 鎌倉事業部内 Kanagawa (JP). 藤木 宏明 (FUJIKI, Hiroaki) [JP/JP]; 〒247-0065 神奈川県 鎌倉市 上町屋 5 2 4 番地 三菱スペース・ソフトウエア株式会社 鎌倉事業部内 Kanagawa (JP). 大山 徹 (OHYAMA, Tohru) [JP/JP]; 〒247-0065 神奈川県 鎌倉市 上町屋 5 2 4 番地 三菱スペース・ソフトウエア株式会社 鎌倉事業部内 Kanagawa (JP). 立野井 亮 (TATENOI, Ryo) [JP/JP]; 〒247-0065 神奈川県 鎌倉市 上町屋 5 2 4 番地 三菱スペース・ソフトウエア株式会社 鎌倉事業部内 Kanagawa (JP).

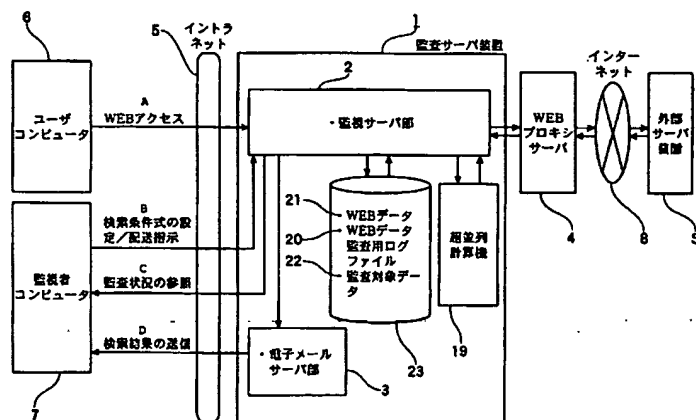
(74) 代理人: 佐野 弘 (SANO, Hiroshi); 〒104-0033 東京都 中央区 新川 2 丁目 3 番 7 号 浪商ビル 3 F Tokyo (JP).

(81) 指定国 (国内): US.

[続葉有]

(54) Title: WEB MONITORING SYSTEM AND METHOD

(54) 発明の名称: WEB 監査システム及び方法



- 6...USER COMPUTER  
7...MONITOR COMPUTER  
A...WEB ACCESS  
B...INSTRUCTION TO SET/TRANSMIT SEARCH CONIDTON FORMULA  
C...REFERENCE TO MONITORING PROCESS  
D...TRANSMISSION OF SEARCH RESULT  
5...INTRANET  
1...MONITORING SERVER DEVICE  
2...MONITOR SERVER UNIT  
21...WEB DATA  
20...WEB DATA MONITORING LOG FILE  
22...DATA MONITORED  
19...MASSIVELY PARALLEL COMPUTER  
3...ELECTRONIC MAIL SERVER UNIT  
4...WEB PROXY SERVER  
8...INTERNET  
5...EXTERNAL SERVER DEVICE

(57) Abstract: A WEB monitoring system has a monitoring server device (1) comprising a WEB access data collecting unit (11) for collecting WEB data acquired from an external WEB site (24) by a user computer (6), an automatically collected WEB data collected automatically by an information collecting robot unit (18) that the monitoring server device (1) has, and provision data provided from the user computer

[続葉有]



(84) 指定国 (広域): ヨーロッパ特許 (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

添付公開書類:

— 国際調査報告書

(6) to the external WEB site (24), a search formula setting unit (12) for setting a search condition formula for extracting improper information from the data being monitored, and a searching unit (13) for executing search by the search condition formula set by the search formula setting unit (12) and classifying the data into at least two kinds according to the degree of match between the data being monitored and the search condition formula. The results of the search by the searching unit (13) can be referred to by a monitor computer (7).

(57) 要約:

監査サーバ装置1には、外部のWEBサイト24からユーザコンピュータ6が入手したWEBデータ及び監査サーバ装置1が有する情報収集ロボット部18によって自動的に収集した自動収集WEBデータ並びにユーザコンピュータ6から外部のWEBサイト24へ提供する提供データを収集するWEBアクセスデータ収集部11と、これらの監査対象データから不適切な情報を抽出するための検索条件式を設定する検索式設定部12と、該検索式設定部12によって設定された検索条件式による検索を実行しこれらの監査対象データと検索条件式との適合度により少なくとも2種以上に分類する検索処理部13と、該検索処理部13による検索結果を監査者コンピュータ7で参照できるようにしたWEB監査システム。

## 明 細 書

### WEB監査システム及び方法

#### 技術分野

この発明は、特定のネットワークを介して外部のWEBサイトにアクセスする場合のアクセス状況の監査に関するものであり、詳しくは、会社又は機関の内部のネットワークから外部のWEBサイトへアクセスして情報を収集又は提供する場合の入手したWEBデータ及び提供する提供データについて、不適切な情報の有無を監査するシステム及びその方法に関するものである。

#### 背景技術

従来、多くの会社又は機関では、内部のネットワークであるイントラネット等から外部のWEBサイトにアクセスする場合には、各種のWEB監査システムが導入されており、それらのシステムを使用して具体的な監査業務が運営されている。その多くは、既に公開されているWEBサイトからの情報を定期的に収集したWEBデータにおいて不適切な情報の検出を行うといった方法を採用している。

しかしながら、実際の監査業務に際しては、検索条件式用キーワードの選定においては、元となる既存文書からこれらのキーワードを監査者が目視により試行錯誤しながら設定しているのが実状である。

また、外部のWEBサイトにアクセスして入手したWEBデータが検索条件式に合致した場合においても、検索条件式に合致したか否かの判定がなく、多くの場合、その内容については監査者の目視による判定に依存しており、監査者の負荷が高いものであった。

## 発明の開示

この発明は、以上のような事情に鑑みて考えられたもので、内部ネットワークから外部への不適切な情報の発信を短時間で検出し、この不適切な情報に迅速な対応措置を講じることができ、検索条件式用キーワードの設定をしなくともよく、かつ検索結果の目視確認における負荷を軽減することのできるWEB監査システム及びその監査方法を提供することを課題としている。

この発明の他の課題は、検索条件式の増加に伴う監査業務の遅延を防止することにある。

かかる課題を解決するために、第1の発明は、特定のネットワークにユーザコンピュータと監査者コンピュータ及び監査サーバ装置とが接続されており、前記ユーザコンピュータが前記ネットワークの外部のWEBサイトにアクセスする場合に該アクセス状況を前記監査サーバ装置が監査するようにした監査システムにおいて、前記監査サーバ装置には、前記外部のWEBサイトから前記ユーザコンピュータが入手したWEBデータ及び前記監査サーバ装置が有する情報収集ロボット部によって自動的に収集した自動収集WEBデータ並びに前記ユーザコンピュータから外部のWEBサイトへ提供する提供データを収集するWEBアクセスデータ収集部と、前記WEBデータ、自動収集WEBデータ及び提供データから不適切な情報を抽出するための検索条件式を設定する検索式設定部と、該検索式設定部によって設定された前記検索条件式による検索を実行し前記WEBデータ、自動収集WEBデータ及び提供データを前記検索条件式との適合度により少なくとも2種以上に分類する検索処理部と、該検索処理部による検索結果を前記監査者コンピュータで参照できるようにする検索結果参照部とからなることを特徴としている。

これによれば、検索結果の参照においては、適合度により目視による確認を行う上での優先度付け並びに絞り込みを行うことが容易にできるため、監査業務の負荷を軽減できる。

第2の発明は、第1の発明の構成に加えて、前記監査サーバ装置には、予め設定した監査対象除外IPアドレス及び監査対象者IPアドレスを参照して前記WEBデータ及び提供データから監査対象データを抽出する監査対象確認部を設けたことを特徴としている。

これによれば、監査対象とする必要のないWEBアクセスを確認しないで済むため、第1の発明の効果に加えて、より監査業務の負荷を軽減できる。

第3の発明は、第1又は2の発明の構成に加えて、前記監査サーバ装置には、予め設定したメールアドレスを有する前記監査者コンピュータに前記検索条件式との適合度により拒絶された前記WEBデータ、自動収集WEBデータ及び提供データに関する検知通知メールを作成する検知通知部と、前記検知通知メールを前記監査者コンピュータに送信する電子メールサーバ部とを設けたことを特徴としている。

これによれば、監査者は自動的に不適切な情報の存在を知ることができるため、第1又は2の発明の効果に加えて、より監査業務の負荷を軽減できる。

第4の発明は、第1乃至3の発明のいずれか1つの構成に加えて、前記監査サーバ装置には、WEBデータ監査用ログを出力するWEBデータ監査用ログ出力部を設けたことを特徴としている。

これによれば、第1乃至3の発明のいずれか1つの効果に加えて、監査者は必要に応じていつでも監査記録を確認することができる。

第5の発明は、第1乃至4の発明のいずれか1つの構成に加えて、前記検索条件式には予め設定したキーワードの他、自然文を入力することができることとし、前記検索式設定部には、自然文が入力された場合に、該自然文を品詞単位に分割し、該品詞単位に分割された言葉において連続する名詞同士を結合する複合語処理を行って複合語を作成すると共に、この作成された複合語を前記検索条件式に加える複合語処理部を有することを特徴としている。

これによれば、第1乃至4の発明のいずれか1つの効果に加えて、自然文

を入力する際に、既存文書からのキーワード選定を行う必要がないため、キーワードの重要語を特に意識することなく、当該既存文書との関連のある電子メールを検知することが可能となる。

第6の発明は、特定のネットワークにユーザコンピュータと監査者コンピュータ及び監査サーバ装置とが接続されており、前記ユーザコンピュータが前記ネットワークの外部のWEBサイトにアクセスする場合に該アクセス状況を前記監査サーバ装置が監査するようにした監査方法において、前記監査者コンピュータから入力されたキーワード又は自然文に基づいて不適切な情報を検索するための検索条件式を作成する検索条件式作成工程と、前記外部のWEBサイトから前記ユーザコンピュータが入手したWEBデータ及び前記監視サーバ装置が有する情報収集ロボット部によって自動的に収集した自動収集WEBデータ並びに前記ユーザコンピュータから外部のWEBサイトへ提供する提供データと前記検索条件式とを計算機に投入して、前記WEBデータ、自動収集WEBデータ及び提供データとからなる監査対象データと前記検索条件式との適合度により不適切な情報を選別する検索処理工程とからなり、前記計算機は複数のプロセッサを内蔵した超並列計算機であって、前記複数のプロセッサ上にそれぞれ独立した複数の前記検索条件式を設定して前記監査対象データを全文検索するようにしたことを特徴としている。

これによれば、検索条件式の設定において自然文を入力する際に、既存文書からのキーワード選定を行う必要がないため、キーワードの重要語を特に意識することなく、当該既存文書との関連のある不適切な情報を検知することが可能となる。

さらに、超並列計算機による全文検索を行うことにより、監査対象データ入手後瞬時に検索結果を得ることが可能であり、WEB監査業務の迅速化が図れる。

### 図面の簡単な説明

第 1 図は、この発明の W E B 監視システムに係る実施の形態の構成を示すブロック図である。

第 2 図は、同実施の形態の監査サーバ部の構成を示すブロック図である。

第 3 図は、同実施の形態の超並列計算機における処理状態を示す模式図である。

第 4 図は、同実施の形態を使用した W E B 監査方法の全体の流れを示すフローチャートである。

第 5 図は、第 2 図における検索条件式の設定において自然文を入力した場合のフローチャートである。

### 発明を実施するための最良の形態

以下、この発明の W E B 監査システムに係る実施の形態について図面に従って説明する。

第 1 図は、この発明の W E B 監査システムに係る実施の形態の構成を示すブロック図である。

第 1 図において、監査サーバ装置 1 は、W E B 監査システムの統括管理を行うコンピュータであり、監査サーバ部 2 及び電子メールサーバ部 3 を有し、外部の W E B プロキシサーバ 4 と接続されている。

監査サーバ装置 1 は、会社又は機関の内部に設けられたイントラネット 5 を介して、クライアント装置であるユーザコンピュータ 6 及び監査者コンピュータ 7 と接続されている。また、監査サーバ装置 1 は、W E B サイトを有する外部のサーバ装置 S と W E B プロキシサーバ 4 を経てインターネット 8 を介して接続されるようになっている。

ユーザは、クライアント装置であるユーザコンピュータ 6 から W E B サイトを有する外部のサーバ装置 S にアクセスする。監査者は、監査用検索条件

式の設定、監査状況の確認を同じくクライアント装置である監査者コンピュータ 7 により行う。

監査サーバ部 2 には、第 2 図に示したように、監査対象確認部 10、WEB アクセスデータ収集部 11、検索式設定部 12、検索処理部 13、検索結果参照部 14、中継部 15、WEB 監査用ログ出力部 16、検知通知部 17 及び情報収集ロボット部 18 を有している。また、監査サーバ部 2 には、図 1 に示したように、超並列計算機 19 と WEB データ監査用ログファイル 20、WEB データ 21 及び監査対象データ 22 等を保存するデータベース部 23 とが接続されている。

データベース部 23 は、第 2 図に示したように、各保存ファイル毎に分割してもよいし、一体ものとしてもよい。そこで、全体構成を示した図 1 ではデータベース部 23 を 1 つのものとして示しているが、より詳細な構成を示した図 2 においては、WEB データ監査用ログファイル 20 と監査対象データ 22 とは別々のデータベースとして表している。ここで、監査対象データ 22 とは、ユーザコンピュータ 6 が外部のサーバ装置 S の WEB サイト 24 から受信した WEB データ 21 と情報収集ロボット部 18 が外部のサーバ装置 S の WEB サイト 24 から自動的に収集してきた自動収集 WEB データ 31 の他、ユーザコンピュータ 6 から外部の WEB サイト 24 へ提供する提供データ（図示せず）とを合わせたものをいい、専用の検索ディレクトリ 25 に保存されるようにしている。

検索条件式の設定においては、キーワードの重要語を特に意識せずに、検索条件に適合する既存文書（過去の監査対象データであってもよい）との関連のある WEB データ 21、自動収集 WEB データ 31 及び提供データを検知することを可能とするため、自動的に既存文書の自然文から検索条件式を生成するようにしている。

検索結果の参照においては、検索条件式との適合度により目視による確認



を行う上での優先度付け並びに絞り込みを行うために、WEBデータ21、自動収集WEBデータ31及び提供データから不適切な情報を選別する基準として、検索条件式との適合度におけるしきい値を2種類設定することにより、「合格」、「不合格」に加えて注意を促すための「注意」の3種類に分類する方法を採用している。

具体的には、しきい値として、最大値を100点とした場合に、合格を0点～10点、注意を11点～90点、不合格を91点～100点と設定することが考えられる。この場合、91点以上が配点されたものが拒絶の対象となる。

なお、合否判定には、上記方法の他に、1つでもキーワードが出現したら不合格とする方式若しくは予め設定した特定のキーワードが合致するものに加点（プラス方向の重み）及び減点（マイナス方向の重み）を付与することにより、これらを合算したものを得点とする方式を組み合わせ使用してもよい。

また、WEBデータ21、自動収集WEBデータ31及び提供データからなる監査対象データ22を入手した後のWEB監査業務の迅速化を図る手段として、複数の検索条件式を事前に複数のプロセッサ上に設定でき、かつインデックス付けを行わないテキストデータを投入することで検索結果を得ることが可能な超並列計算機19を採用している。これにより、新規な監査対象データ22を受信するとすぐに超並列計算機19上で検索条件式と入力された監査対象データ22との照合を行うことによる全文検索を行えば、監査対象データ22の受信から検索結果を得るまでの時間を短縮することが可能である。

超並列計算機19は、第3図の模式図に示したように、数千～数万のプロセッサ26, 26, 26, ...（以下、これらをまとめてパイプラインという。）を内蔵することにより、このパイプライン27に複数の異なった検

索条件式 28 を同時に設定可能としている。そして、これら大量のプロセッサ 26, 26, 26, . . . を同時に動作させることによって、新たな監査対象データ 22 が送信されてくると、パイプライン 27 に監査対象データ 22 を送出し複数の異なった検索条件式 28 と監査対象データ 22 とのマッチングを行う全文検索を実行する。マッチングの結果、検索条件式 28 に合致する監査対象データ 22 が見つかったら、その監査対象データ 22 がヒットした（「不合格」又は「注意」）とみなす機能を有している。

ここで、超並列計算機 19 は、全文検索エンジン（例えば、Parace l 社製、FDF（登録商標）4T Text Finder）のような機器が望ましいが、これと同等の機能及び性能を有するワークステーション等の機器でもよい。

以下、この発明に係るメール監査システムの実施の形態による監査手順を第 4 図及び第 5 図に従って説明する。

第 4 図は、WEB 監査方法の全体の流れを示すフローチャートである。

まず、イントラネット 5 の内部のユーザが、クライアント装置であるユーザコンピュータ 6 から目的とする WEB サイト 24 を有する外部のサーバ装置 S の URL を指定して WEB サイト 24 にアクセスすると（ステップ S 1）、ユーザには意識されずに、その WEB アクセスを監査サーバ部 2 が受信し当該ユーザの IP アドレスを入手して、監査対象確認部 10 で監査除外アドレス定義ファイル 31 及び監査対象者定義ファイル 32 に保存されている監査対象除外者 IP アドレス、すべて監査対象としないとするフラグ及び監査対象者の IP アドレス等を参照して、その WEB アクセスが監査対象か否かを確認する（ステップ S 2）。その結果、監査対象となった WEB アクセスについてその WEB データ 21 を WEB アクセスデータ収集部 11 が収集する（ステップ S 3）。

また、監査サーバ部 2 の情報収集ロボット部 18 が外部のサーバ装置 S の

WEBサイト24から自動的に収集した場合には（ステップS4）、その自動収集WEBデータ31はそのままWEBアクセスデータ収集部11にて収集される（ステップS3）。

監査対象となったWEBアクセスについては、そのアクセス内容からテキスト部分を抽出し、監査対象データ22としてデータベース部としての検索ディレクトリ25に登録する。また、情報収集ロボット部18により自動的に収集された自動収集WEBデータ31については、その本文の内容をそのまま監査対象データ22として同じくデータベース部としての検索ディレクトリ25に登録する。

監査者は、監査用検索条件式28のためのキーワード又は既存文書（過去の監査対象データであってもよい）の自然文を監査者コンピュータ7から入力する（ステップS5）。すると、その指示に従い検索式設定部12で検索条件式28を生成し、この検索条件式28が検索処理部13へと送られる。

検索条件式28は、複数のキーワードをAND、OR、NOT等の組み合わせたものとなるが、自然文を入力した場合には、自然文からキーワードとなる言葉を自動的に切り出されるようになっている。なお、検索条件式28の自動生成の詳細については後述する。

検索条件式28が与えられると、受信した監査対象データ22がその検索条件式28と合致するかどうか検索処理部13にて検索を開始する（ステップS6）。

検索結果は、監査者コンピュータ7の検索結果参照部14としてのモニター等（図示せず）に表示する。具体的には、検索結果は、検索条件式28との適合度を予め指定された2つのしきい値と比較判断し、「合格」、「不合格」、「注意」の3種類に分類される。

そこで、検索の結果が合格か否かを確認し（ステップS7）、「合格」となった「合格」WEBアクセスは、監査処理を終了する。「合格」でないと

された「注意」又は「不合格」のWEBアクセスについては、その結果のメール送信要望があるか否かを確認し（ステップS8）、要望がある場合には電子メールサーバ部3を介して検索結果を指定されたメールアドレス（例えば、監査者コンピュータ7）へ送信し（ステップS9）、検索結果をそのメールアドレスのコンピュータの表示部に表示すると共に（ステップS10）、監査対象データ22をデータベース部20に保存して監査処理を終了する。

「合格」でないとされた「注意」又は「不合格」のWEBアクセスについては、中継部15で作成された警告メッセージ34が当該WEBアクセスを行ったユーザーコンピュータ6へ送信される。

また、「合格」でないとされた「注意」又は「不合格」のWEBアクセスに関する監査対象データ22を指定されたメールアドレスに送信するに当たっては、検知通知部17にて通知先アドレス定義29を参照して、正しい通知先のコンピュータのIPアドレスのメールアドレスへ送信される。このとき、検知通知部17では、並行して検知通知メール30が作成され、この検知通知メール30が電子メールサーバ部3を介して監査者コンピュータ7に送信される。ここで、電子メールサーバ部3は、図2に示したように、メール受信サーバ部3aとメール転送サーバ部3bとで構成することで、ユーザーコンピュータ6が常時接続され稼動していなくともユーザーコンピュータ6が接続したときに自分宛の電子メールが受信できるようにしている。

ユーザーコンピュータ6から受信したWEBアクセスについての監査状況は、WEBデータ監査用ログ出力部16によってログとして出力される。このログには、WEBアクセス先であるWEBサイト24のURL、発信元のIPアドレス及び合格、不合格、注意の別等が記載され、WEBデータ監査用ログとしてWEBデータ監査用ログファイル20に保存される。また、WEBデータ監査用ログには、監査の結果として条件に合致したか否かの判定及び検索条件式との適合度が含まれる。

次に、自然文による検索条件式の自動生成の詳細を、第5図に従って説明する。

第5図は、検索条件式の設定において自然文を入力した場合のフローチャートである。

まず、監査者が入力した自然文を形態素解析により品詞単位に分割し、品詞情報を取得する（ステップS21）。次に、品詞単位に分割した言葉を、例えば、連続する名詞は結合させる等の複合語処理することにより、複合語を作成する（ステップS22）。この複合語を検索キーワードとして採用することで、言葉の分割による表現の抽象化等を回避することができ、条件を適確に表す言葉を抽出することができる。

次に、データベース部23に保存されている過去の監査対象データ22を統計処理した統計情報をもとにして、品詞情報、言葉が現れる監査対象データ数、言葉の一般性等からキーワードを選定する（ステップS23）。

ここで、品詞情報、言葉が現れる監査対象データ数、言葉の一般性の選定条件を詳細に説明する。

まず、品詞情報から、複合語、名詞、未定義語等、特徴を表す品詞の言葉を抽出する。抽出対象とする品詞は、上記の複合語、名詞、未定義語に限られるものではなく、配信対象となる監査対象データ22毎に特徴を表すと予測される品詞を指定する。品詞単位に分割された後、場合によっては、必ず削除する言葉等をリストとして保有し、抽出対象品詞であっても、その言葉を削除する機能を追加してもよい。

品詞情報により、絞り込まれた言葉に対し、言葉の出現回数や、言葉の一般性を表す重要度等からその言葉に対する重みを算出する。

言葉の出現回数は、出現回数が多い言葉は、その記事の概念を表す言葉と考えられる。

言葉の一般性を表す重要度は例えば、様々な監査対象データ22を格納し

たデータベース（検索ディレクトリ 25）における、その言葉が出現する記事の割合を算出する。その割合が多ければ様々な記事に出現する言葉であるため、重要度が小さくなり、割合が少なければ、ある特定の記事にのみ現れる言葉として、重要度が大きくなる。

出現頻度が多い言葉として、一般的な言葉も考えられるが、そのような言葉は、言葉の重要度が小さく、その言葉に付与される重みは、小さくなるため、検索結果に悪影響を与えることは少ない。

キーワードが選定されたら、そのキーワードをもとに、指定した既存文書等に関連した監査対象データ 22 を検索するための（キーワード、重み、スコアリング方式等からなる）検索条件式を自動生成する（ステップ S 24）。スコアリング方式としては、監査対象データ 22 内に一度でも出現すればスコアを計上する方式や、監査対象データ 22 内で、出現する度にスコアを計上する方法等があげられる。

上述した実施の形態では、ユーザコンピュータ 6 及び監査者コンピュータ 7 と監査サーバ装置 1 との接続としてイントラネット 5 をその接続手段としているが、WEB プロキシサーバ 4 に接続可能な他のネットワーク等の接続でもよい。

監査者コンピュータ 7 は、監査サーバ装置 1 と兼用してもよい。

監査サーバ装置 1 に含まれている監査サーバ部 2、電子メールサーバ部 3 は、それぞれ独立した構成であってもよい。WEB プロキシサーバ 4 は監査サーバ装置 1 と別体としているが監査サーバ装置 1 に組み込んで WEB プロキシサーバ部を構成するようにしてもよい。つまり、監視サーバ部 2、電子メールサーバ部 3 及び WEB プロキシサーバ 4 のいずれか若しくはすべてが監査サーバ部 2 と配置的に同じ場所に設置されていなくても、ネットワーク等で接続されている形態であればよい。

また、ユーザコンピュータ 6 としては、インターネット又は電子メールサ

ービスが受けられる携帯電話等を使用してもよい。

#### 産業上の利用可能性

この発明は、会社又は機関の内部のネットワークから外部のWEBサイトへアクセスして情報を収集又は提供する場合の入手したWEBデータ及び提供する提供データについて、不適切な情報の有無を監査するシステム及びその方法に関するものであり、内部ネットワークから外部への不適切な情報の発信を短時間で検出し、この不適切な情報に迅速な対応措置を講じることができ、検索条件式用キーワードの設定をしなくともよく、かつ検索結果の目視確認における負荷を軽減することのできるため、WEB監査システムのシステム管理者の監査業務の遅延を防止に有用である。

### 請求の範囲

1. 特定のネットワークにユーザコンピュータと監査者コンピュータ及び監査サーバ装置とが接続されており、前記ユーザコンピュータが前記ネットワークの外部のWEBサイトにアクセスする場合に該アクセス状況を前記監査サーバ装置が監査するようにした監査システムにおいて、前記監査サーバ装置には、前記外部のWEBサイトから前記ユーザコンピュータが入手したWEBデータ及び前記監査サーバ装置が有する情報収集ロボット部によって自動的に収集した自動収集WEBデータ並びに前記ユーザコンピュータから外部のWEBサイトへ提供する提供データを収集するWEBアクセスデータ収集部と、前記WEBデータ、自動収集WEBデータ及び提供データから不適切な情報を抽出するための検索条件式を設定する検索式設定部と、該検索式設定部によって設定された前記検索条件式による検索を実行し前記WEBデータ、自動収集WEBデータ及び提供データを前記検索条件式との適合度により少なくとも2種以上に分類する検索処理部と、該検索処理部による検索結果を前記監査者コンピュータで参照できるようにする検索結果参照部とからなることを特徴とするWEB監査システム。

2. 前記監査サーバ装置には、予め設定した監査対象除外IPアドレス及び監査対象者IPアドレスを参照して前記WEBデータ及び提供データから監査対象データを抽出する監査対象確認部を設けたことを特徴とする請求項1に記載のWEB監査システム。

3. 前記監査サーバ装置には、予め設定したメールアドレスを有する前記監査者コンピュータに前記検索条件式との適合度により拒絶された前記WEBデータ、自動収集WEBデータ及び提供データに関する検知通知メールを作成する検知通知部と、前記検知通知メールを前記監査者コンピュータに送信する電子メールサーバ部とを設けたことを特徴とする請求項1又は2に記載



のWEB監査システム。

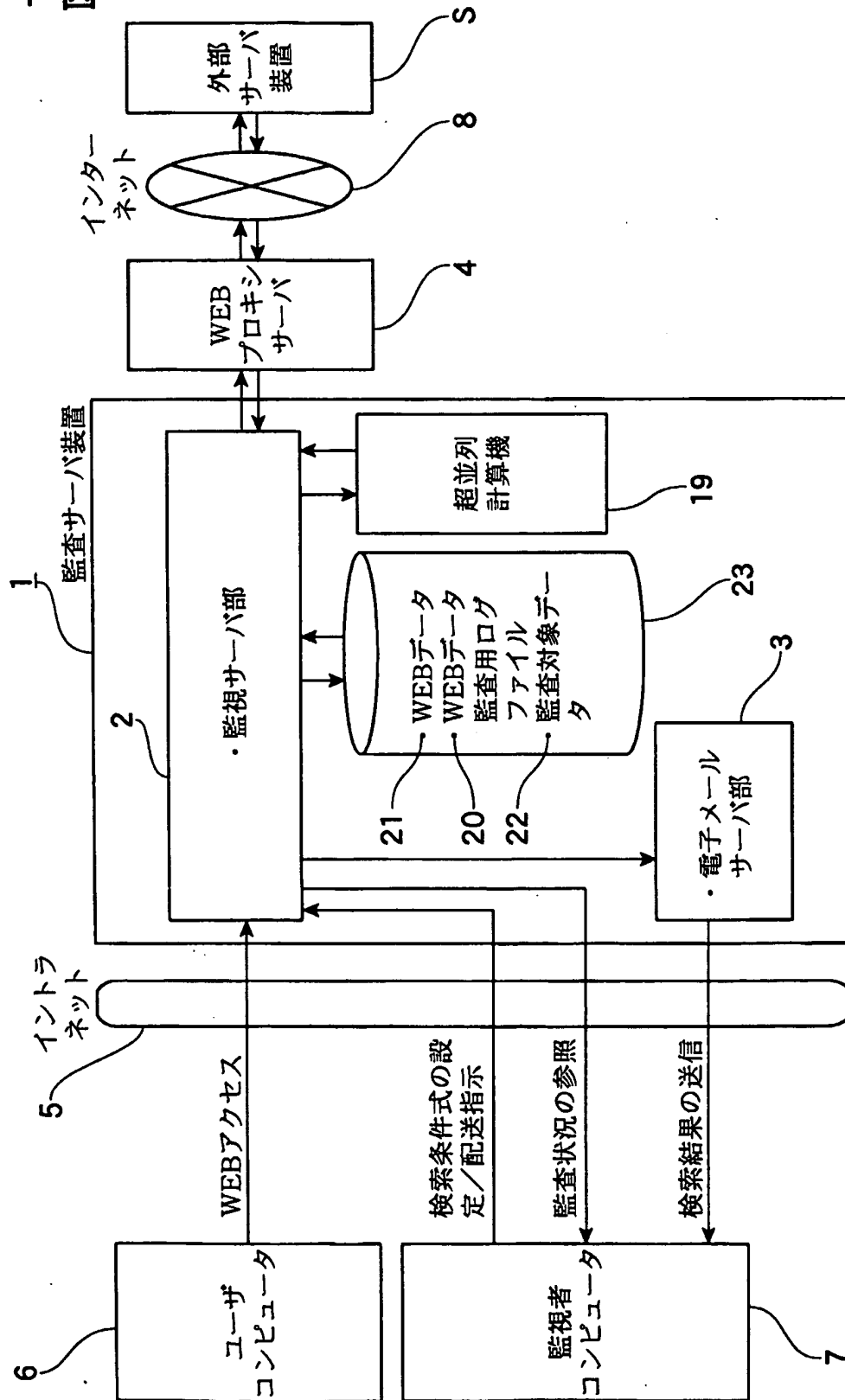
4. 前記監査サーバ装置には、WEBデータ監査用ログを出力するWEBデータ監査用ログ出力部を設けたことを特徴とする請求項1乃至3にいずれか1つに記載のWEB監査システム。

5. 前記検索条件式には予め設定したキーワードの他、自然文を入力することができることとし、前記検索式設定部には、自然文が入力された場合に、該自然文を品詞単位に分割し、該品詞単位に分割された言葉において連続する名詞同士を結合する複合語処理を行って複合語を作成すると共に、この作成された複合語を前記検索条件式に加える複合語処理部を有することを特徴とする請求項1乃至4のいずれか1つに記載のWEB監査システム。

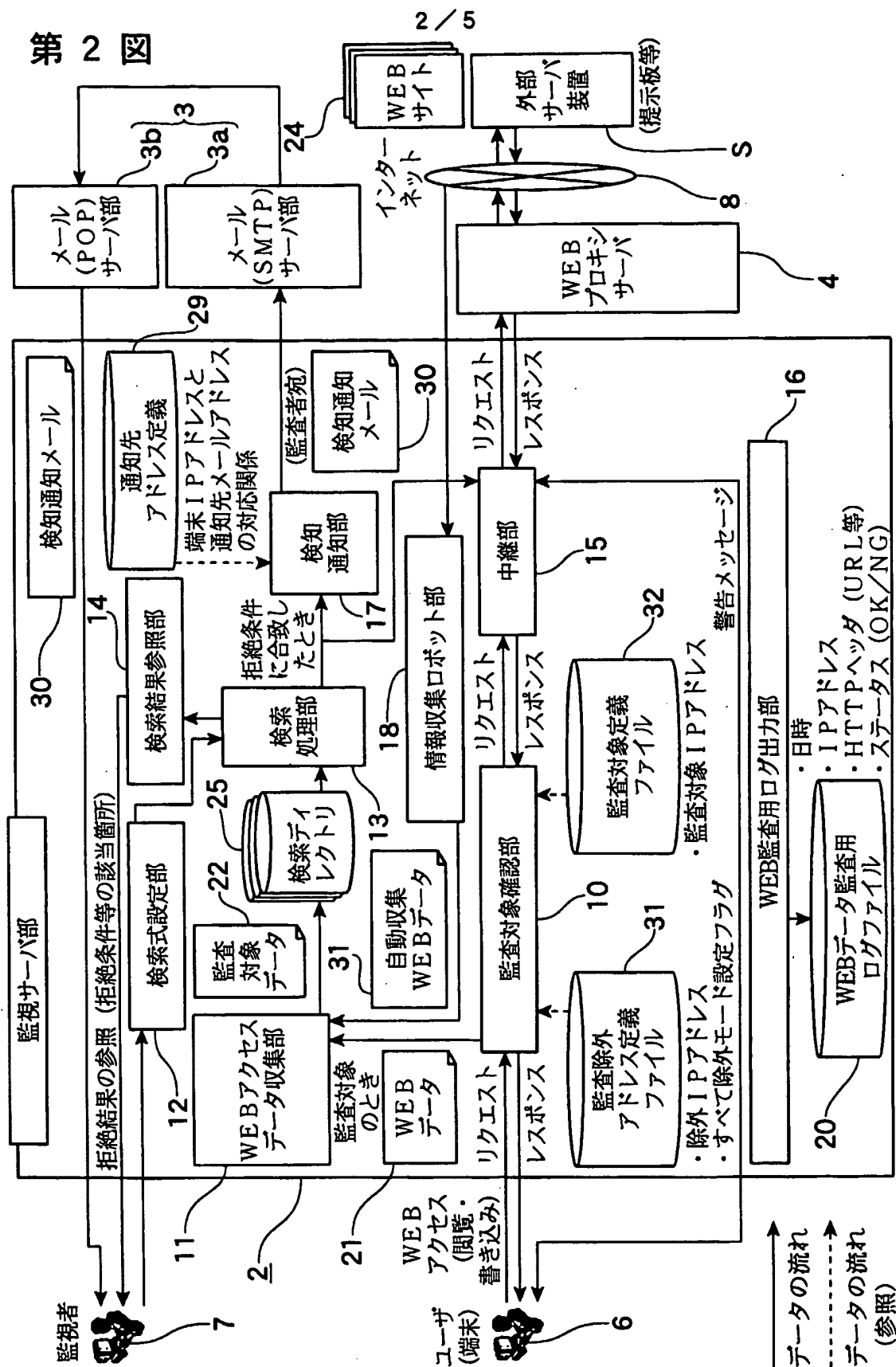
6. 特定のネットワークにユーザコンピュータと監査者コンピュータ及び監査サーバ装置とが接続されており、前記ユーザコンピュータが前記ネットワークの外部のWEBサイトにアクセスする場合に該アクセス状況を前記監査サーバ装置が監査するようにした監査方法において、前記監査者コンピュータから入力されたキーワード又は自然文に基づいて不適切な情報を検索するための検索条件式を作成する検索条件式作成工程と、前記外部のWEBサイトから前記ユーザコンピュータが入手したWEBデータ及び前記監査サーバ装置が有する情報収集ロボット部によって自動的に収集した自動収集WEBデータ並びに前記ユーザコンピュータから外部のWEBサイトへ提供する提供データと前記検索条件式とを計算機に投入して、前記WEBデータ、自動収集WEBデータ及び提供データとからなる監査対象データと前記検索条件式との適合度により不適切な情報を選別する検索処理工程とからなり、前記計算機は複数のプロセッサを内蔵した超並列計算機であって、前記複数のプロセッサ上にそれぞれ独立した複数の前記検索条件式を設定して前記監査対象データを全文検索するようにしたことを特徴とするWEB監査方法。

## 第 1 図

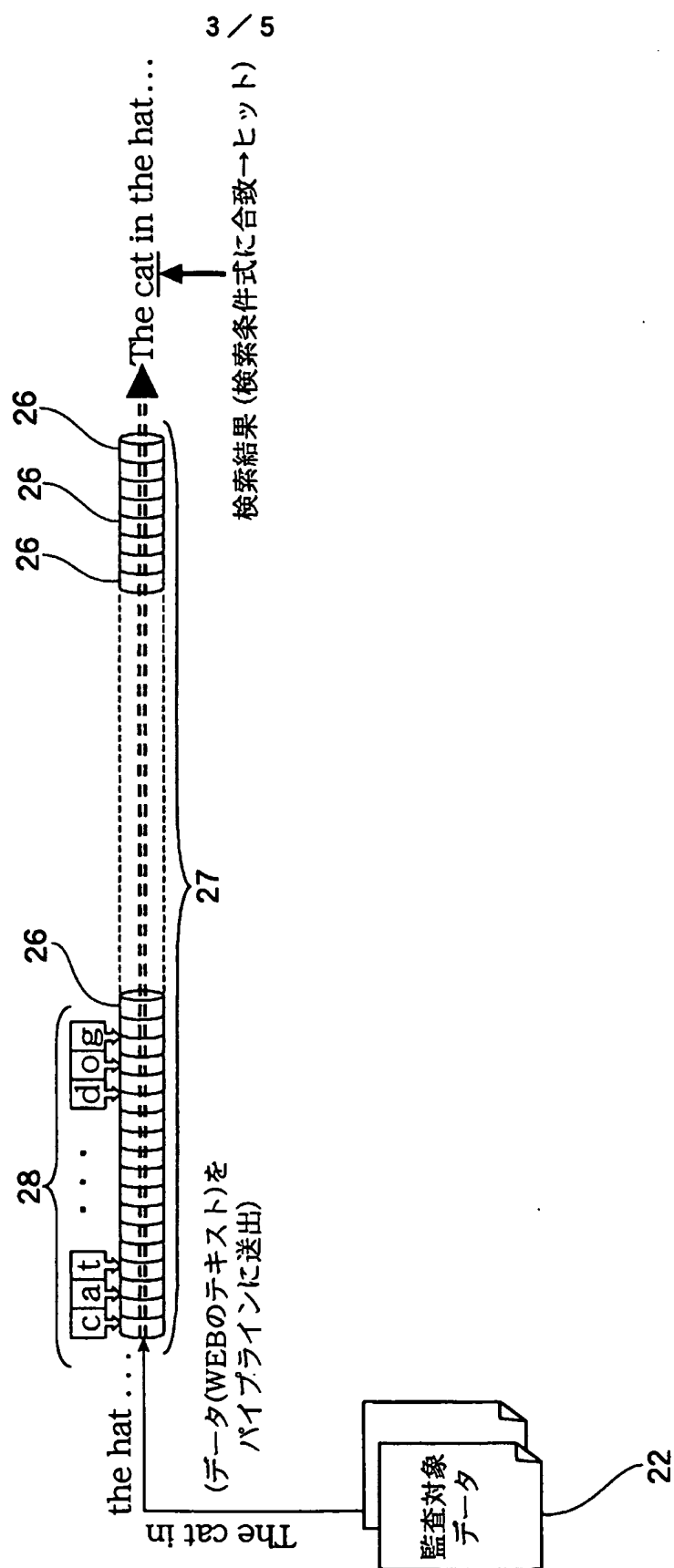
1 / 5



第 2 図

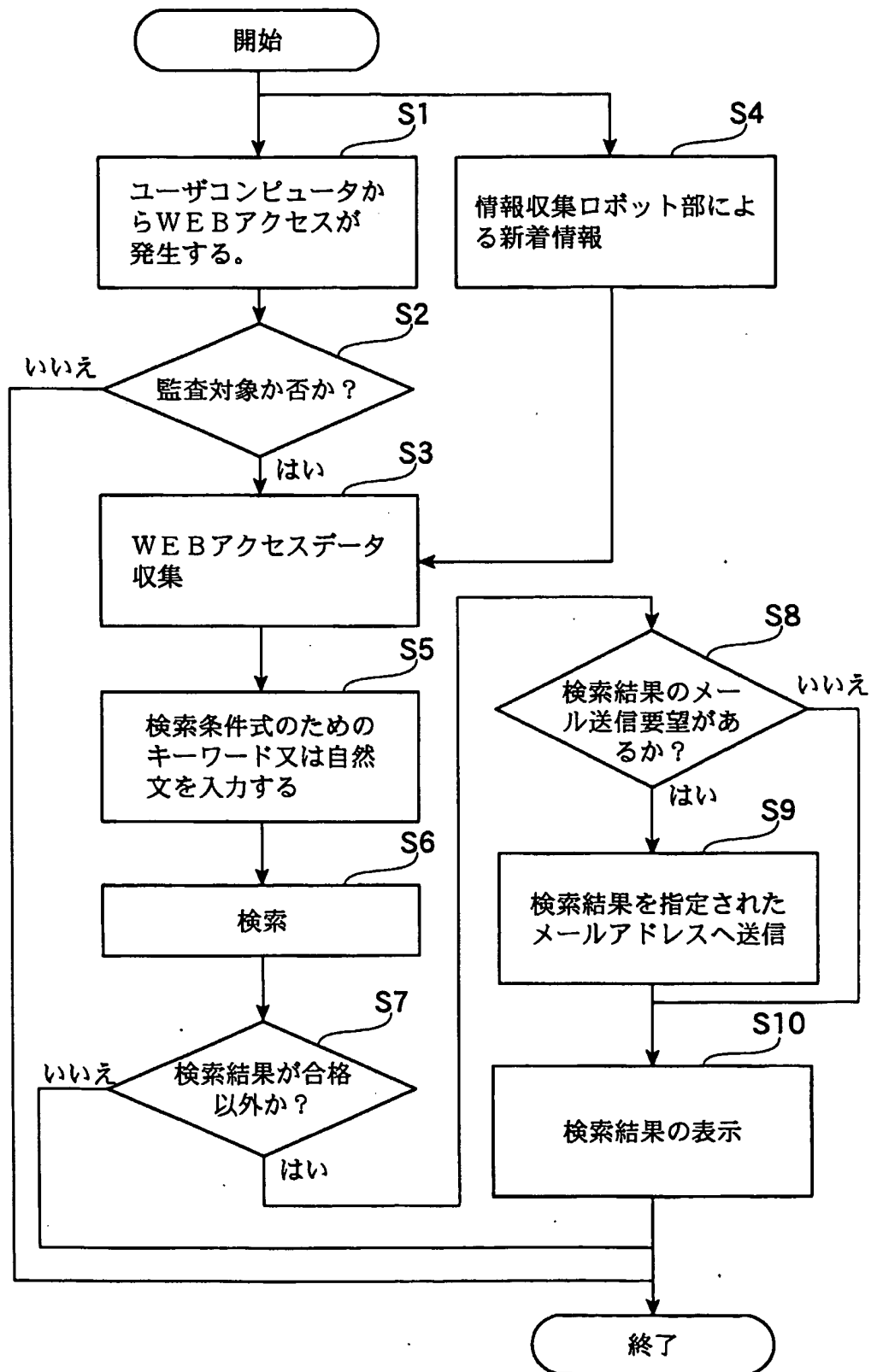


第 3 図



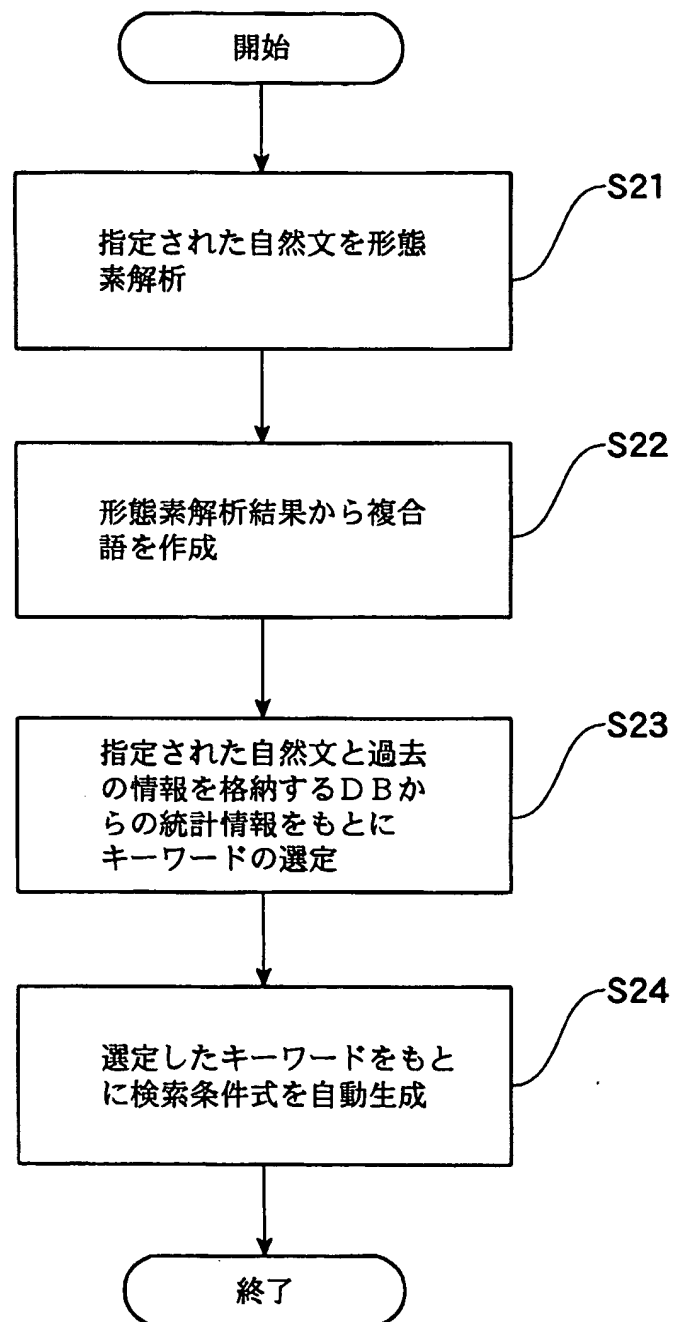
4 / 5

## 第 4 図



5 / 5

## 第 5 図



# INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP02/02639

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl<sup>7</sup> G06F17/30, 15/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl<sup>7</sup> G06F17/30, 15/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2002
Kokai Jitsuyo Shinan Koho	1971-2002	Jitsuyo Shinan Toroku Koho	1996-2002

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JICST FILE (JOIS)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 11-219363 A (Toshiba Corp.), 10 August, 1999 (10.08.99), Par. Nos. [0051] to [0054]; Figs, 12 to 13 (Family: none)	1-6
Y	Yasumori TAKIZAWA, "Internet no 'Monban' Filtering Tool no Koyo", Nikkei Communication No.256, Nikkei Business Publications, Inc., 20 October, 1997 (20.10.97), pages 136 to 139, particularly, page 138, right column, line 8 to page 139, left column, line 11	1-6
Y	Hidenori KAWAMATA, Jyujitsusuru Kanshi - Seigenyo Soft, Nikkei Computer No.469, Nikkei Business Publications, Inc., 10 May, 1999 (10.05.99), pages 87 to 91, particularly, page 91, left column, lines 18 to 31	2-5

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search  
11 April, 2002 (11.04.02)

Date of mailing of the international search report  
23 April, 2002 (23.04.02)

Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP02/02639

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Noriharu TANAKA, "Contents Filter Kino o Motta Firewall", LAN TIMES Vol.8, No.12, Softbank Corp., 01 December, 1998 (01.12.98), pages 128 to 130, particularly, page 130, left column, line 27 to middle column, line 12	3-5
Y	Kenji HONMA, "Anatano Pasocon ha Kanshi Sareteiru", Nikkei Pasocon No.377, Nikkei Business Publications, Inc., 22 January, 2001 (22.01.01), pages 138 to 145, particularly, page 140, right column, line 6 to 8	4-5
Y	JP 11-203309 A (NTT Data Corp.), 30 July, 1999 (30.07.99), Abstract; Par. No. [0031] (Family: none)	5
A	Akio KOKUBUN et al., "Internet ni Okeru Contents Rating to Filtering, Johoshori Vol.40, No.1, 15 January, 1999 (15.01.99), IPSJ: Information Processing Society of Japan, pages 57 to 61, particularly, page 60, right column, line 11 to page 61, left column, line 20	1-6
A	Yasumori TAKIZAWA, "Donyu Kiun Takamaru Web Filtering Shiyo Web Surfing o Gekitai Access Kinshi Settei ga Jyunan ni", Nikkei Communication No.293, Nikkei Business Publications, Inc., 03 May, 1999 (03.05.99), pages 44 to 45	1-6



## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl. G06F17/30, 15/00

## B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl. G06F17/30, 15/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年  
 日本国公開実用新案公報 1971-2002年  
 日本国登録実用新案公報 1994-2002年  
 日本国実用新案登録公報 1996-2002年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JICSTファイル (JOIS)

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP 11-219363 A (株式会社東芝) 1999.08.10, 第51段落-第54段落, 第12-13図 (ファミリーなし)	1-6
Y	滝沢 泰盛, インターネットの「門番」フィルタリング・ツールの効用, 日経コミュニケーション第256号, 日経BP社, 1997.10.20, pp. 136-139, 特に第138ページ右欄第8行目-第139ページ左欄第11行目	1-6

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」口頭による開示、使用、展示等に関する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&amp;」同一パテントファミリー文献

国際調査を完了した日

11.04.02

国際調査報告の発送日

23.04.02

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

平井 誠

5M

9071

電話番号 03-3581-1101 内線 3597

C (続き). 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	川又 英紀, 充実する監視・制限用ソフト, 日経コンピュータ No. 469, 日経BP社, 1999. 05. 10, pp. 87-91, 特に第91ページ左欄第18-31行目	2-5
Y	田中 紀治, コンテンツフィルタ機能を持ったファイアウォー ル, LAN TIMES 第8巻第12号, ソフトバンク株式 会社, 1998. 12. 01, pp. 128-130, 特に第 130ページ左欄第27行目-同中欄第12行目	3-5
Y	本間 健司, あなたのパソコンは監視されている, 日経パソコ ン第377号, 日経BP社, 2001. 01. 22, pp. 138-145, 特に第140ページ右欄第6行目- 第8行目	4-5
Y	JP 11-203309 A (株式会社エヌ・ティ・ティ・データ) 1999. 07. 30, 要約, 第31段落 (ファミリーなし)	5
A	国分男ほか, インターネットにおけるコンテンツ・レイティン グとフィルタリング, 情報処理第40巻第1号, 1999. 01. 15, 社団法人情報処理学会, pp. 57-61, 特に第60ページ右欄第11行目-第61頁左 欄第20行目	1-6
A	滝沢泰盛, 導入機運高まるWebフィルタリング 専用Webサ ーフィンを撃退アクセス禁止設定が柔軟に, 日経コミュニケーシ ョン第293号, 日経BP社, 1999. 05. 03 pp. 44-45	1-6